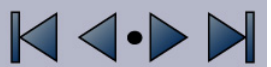
Developing a Navigation and Visualization System for Signaling Pathways

Gilberto R. Sambrano
University of California San Francisco, San Francisco CA

Abstract: Pathway maps are a useful tool for visualizing the complex set of interactions that transmit signals throughout a cell. However, visual representation of a process that is spatially and temporally dynamic and includes hundreds to thousands of distinct molecules is not trivial. We are just beginning to appreciate the tremendous complexity of signaling pathways and to realize that drawings cannot possibly capture all of the facets. This communication is an attempt to consider some of the challenges in visualizing and navigating through complex pathways. The goal is to develop practical and comprehensible maps for both communication and analysis. A system of mapping designations and conventions is proposed that permits the display of whole cellular networks as well as simple signaling paths. Navigation through complex pathways is accomplished by creating molecular designations that can be scaled to different fields of view. The system presented here is a work in progress meant for further exploration and advancement by interested scientists and bioinformaticists.

Published: June 5, 2003
Updated: February 21, 2005

June 5, 2003
Volume 1
DA0009



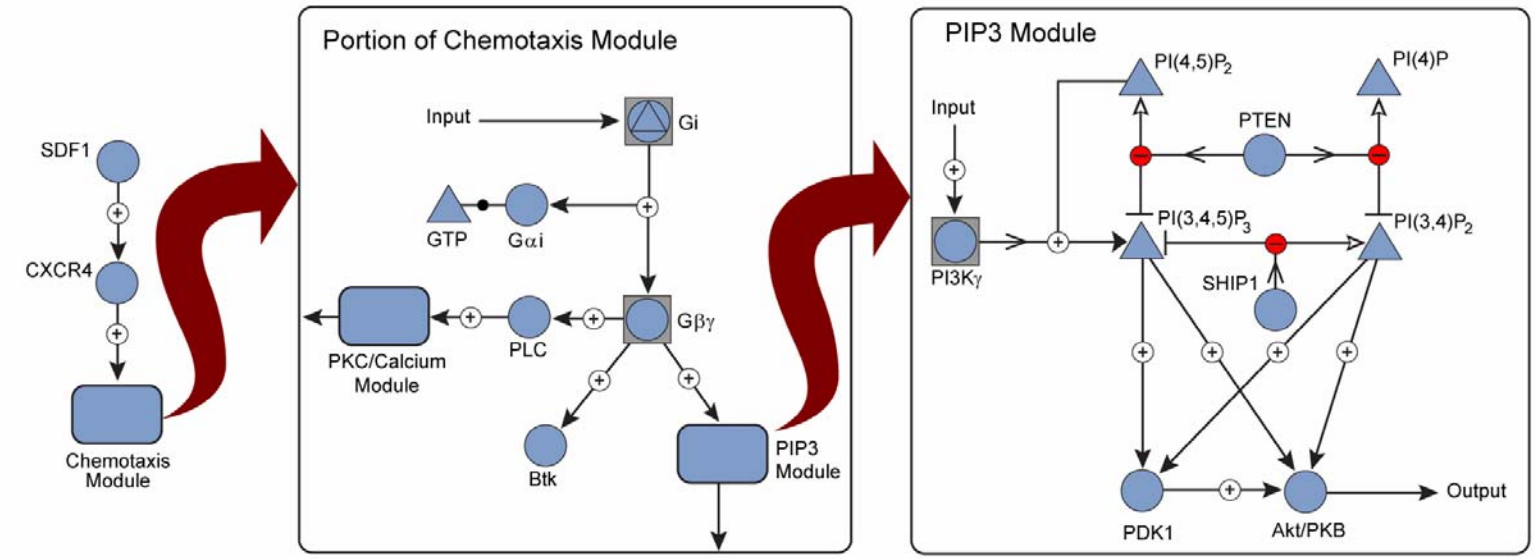
Introduction

Having acquired greater understanding of intracellular signaling mechanisms in recent years, our perception of signaling pathways has evolved from thinking of simple linear paths to conceptualizing more complex signaling networks. Consequently, our need for better visualization tools to navigate through these complex networks has expanded. In addition to tools that organize, parse, or display pathway data is a need to develop a meaningful and practical diagrammatic language for visual representation and navigation. Others have published proposals for diagramming conventions that are useful but do not take us beyond the static, two-dimensional wall map (1,2). Simple diagramming tools that produce connections among signaling molecules are usually sufficient for scientists focused on small areas of any given signaling network. However, as we begin to delve deeper into the specificity of signaling, pathway cross talk, and dynamic aspects of information flow, the drafting of signaling pathways becomes more challenging.

As the Alliance for Cellular Signaling (AfCS) continues its effort to assemble a cellular signaling network, we also intend to develop tools that will facilitate its visualization and analysis. To that end, this communication presents a set of simple diagrammatic and layout guidelines for creating and visualizing signaling pathways. The mapping conventions presented allow one to represent pathways that range from simple linear paths to whole cell signaling networks.

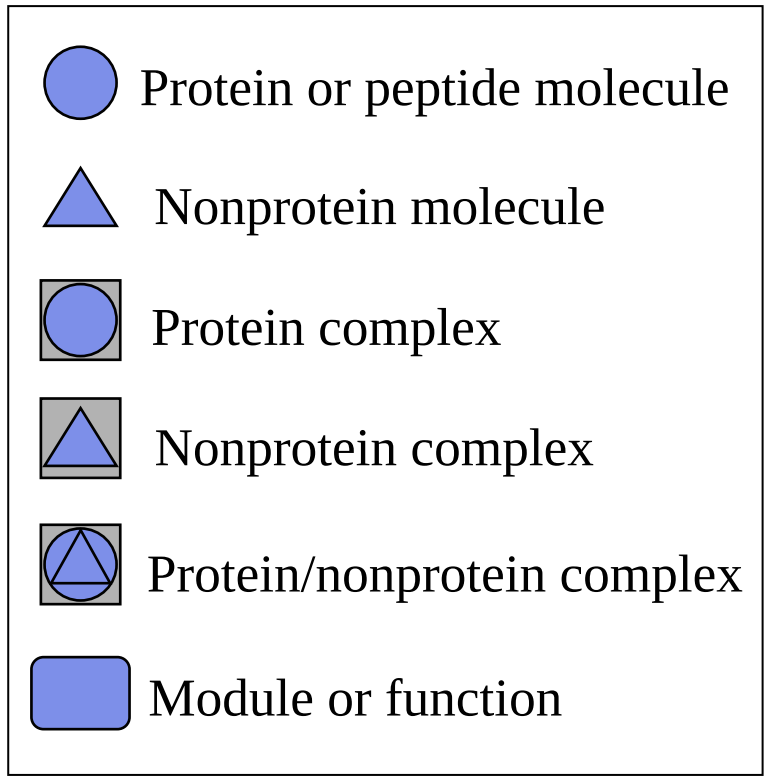
Concept for Creating Navigational Maps

The central concept for creating navigational maps is the ability to scale the field of view. We are quite familiar with its application to geographical maps. For example, we do not typically see street names when viewing a world map. Scalability permits one to navigate from the perspective of a broad cellular network down to the fine details of a specific pathway without being overwhelmed by the multitude of connections. However, since signaling maps are much more dynamic than geographical maps, one must be careful to maintain the integrity of the connections as one moves from one scaled view to another.



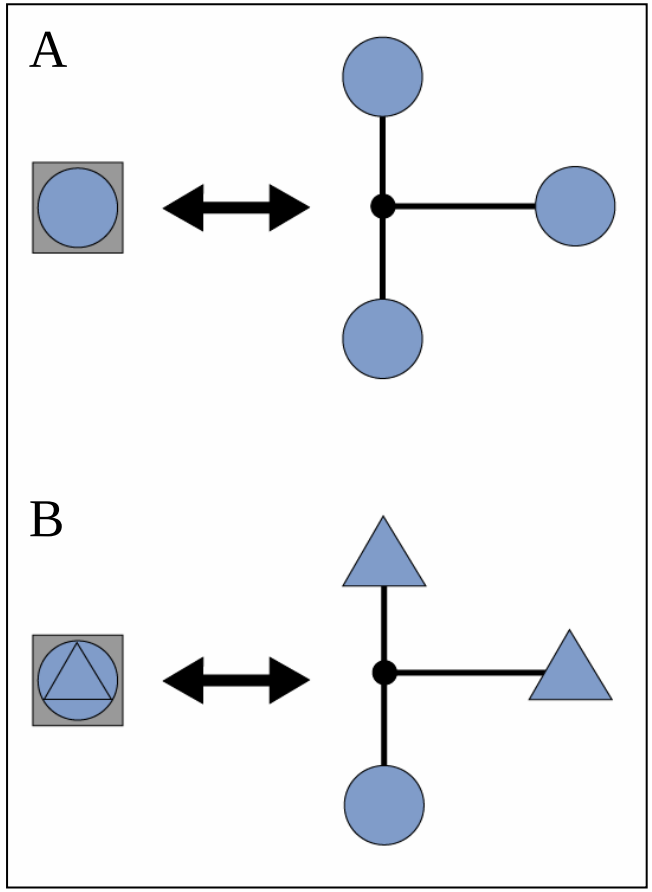
Molecule Designations for Navigation

The shapes shown on this page were designed for navigational maps and are meant to represent molecules as single entities, as complexes, or as components of a module. Symbols distinguish protein and nonprotein molecules as circles and triangles, respectively, while complexes of these are indicated within a gray box. A module could presumably represent a group of molecules and/or molecular complexes that function together or interact. However, for the purpose of pathway navigation, a module can more simply be considered a compartment that contains molecules grouped together for visual convenience. Modules may be assigned names that best describe their contents or general function.



Viewing a Molecular Complex

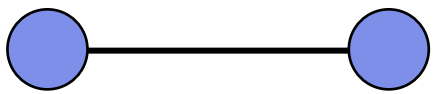
Complexes, whether composed of protein or nonprotein molecules can be represented in several ways. The simplest notation is using the gray box that indicates the general composition of the complex with the protein and/or nonprotein symbols. These complex symbols might be expanded to reveal the specific components that make up the complex. In the examples shown, a protein complex (A) reveals that it is made up of three individual proteins (perhaps a trimer), while a mixed complex (B) reveals a protein and two nonprotein molecules (perhaps a protein bound to two calcium ions). Note that a protein that has been covalently modified with a nonprotein molecule such as a phosphate group (phosphorylation event) could be considered a single entity and thus be represented by a circle. The black node at the center of the three molecules in each example represents the complex formed by the linked molecules.



Making Connections: Edges and Arrows

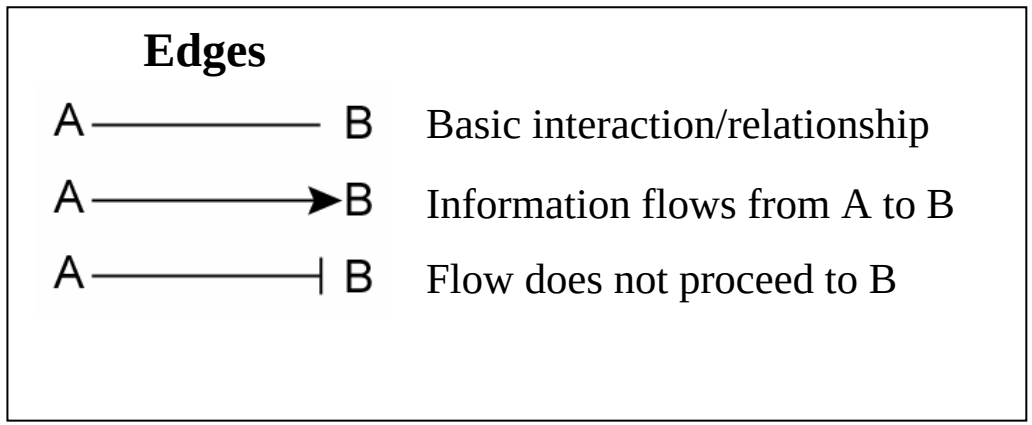
The simplest way to show that a relationship exists between any two signaling molecules is by drawing a line (edge) that connects them. Of course, this is also the least informative representation, as we know that there are many ways in which two or more molecules may interact. The connection is nevertheless valid, and so we might consider a plain line to be the most basic method of indicating an interaction (direct or indirect).

Arrowheads are also often found within signaling pathway maps, but their meaning is nearly as vague as that of the plain line. In many cases, there is an implied positive control on the molecule that the arrow points to. In other cases, the arrow represents a chemical transformation or perhaps a binding interaction. However, for the purpose of pathway navigation, an arrow might best serve to point out the direction of information flow. The transmission of a signal from a specific input to a specific cellular response follows a series of temporally and spatially organized events that are important to highlight. It is, after all, this transmission from one molecule to the next that we are often attempting to trace.



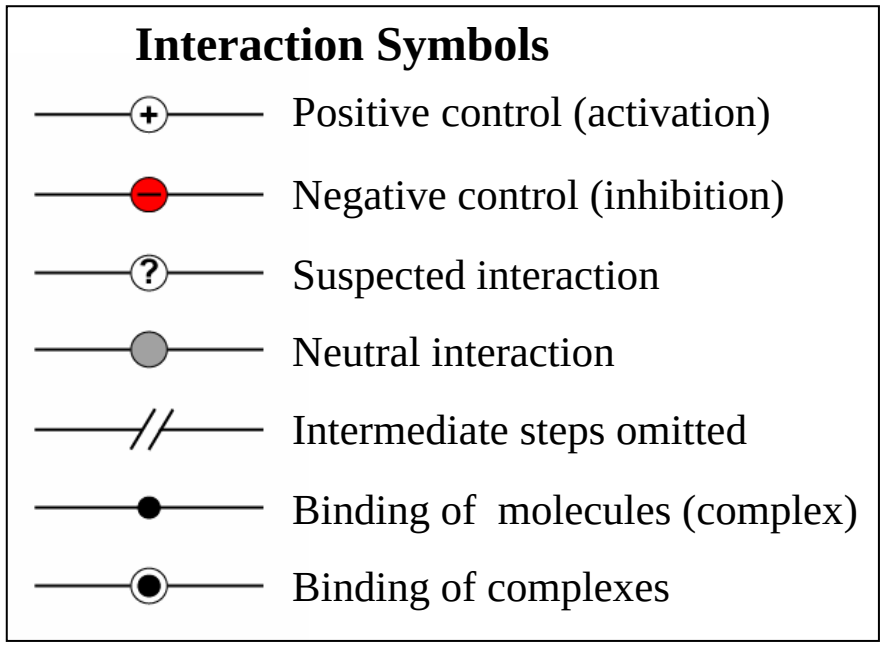
Making Connections: Tracing a Signal

The edges shown below are used to trace the flow of signaling information along a pathway. When the direction of flow is known, an arrow points to the molecule to which information is advanced. In many cases, this will be the product of a regulated reaction or the active state of a molecule. A terminal endpoint indicates that information flow does not proceed to the target molecule, typically because it is inhibited. The joining of two molecules by an edge indicates the existence of a relationship but does not necessarily mean that a direct interaction takes place. That is, several intermediate steps may exist between the molecule transmitting a signal and the one receiving it.

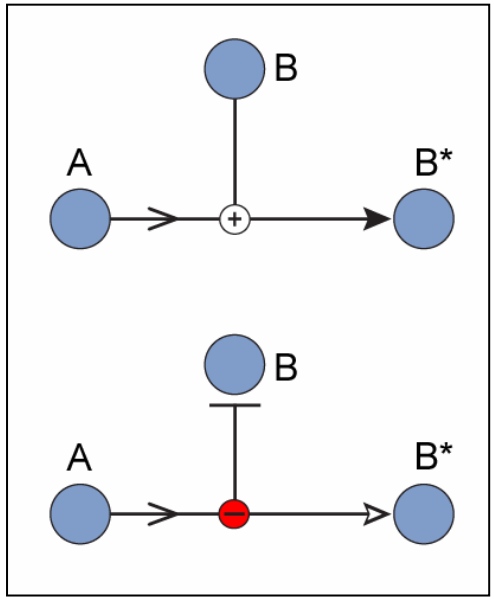


Making Connections: Assigning Meaning to Edges

Once a connection between two molecules has been indicated by an edge, we may also want to show what kind of interaction occurs, if known. There are two general categories that can be used to describe the action of one molecule upon another without knowing specific details of the interaction, positive or negative control. A signaling molecule may increase, induce, or promote the activity of another, and the relationship between them might be defined with a positive symbol. A signaling molecule might also decrease, inhibit, or prevent the activity of another, and their relationship might be defined with a negative symbol. Some interactions are simply suspected, while the outcome of other proven interactions are uncertain (e.g., yeast two-hybrid interactions). The symbols shown represent each of these possible categories. Note that edges identify participants and indicate direction, while symbols define the type of relationship that exists between the molecules.

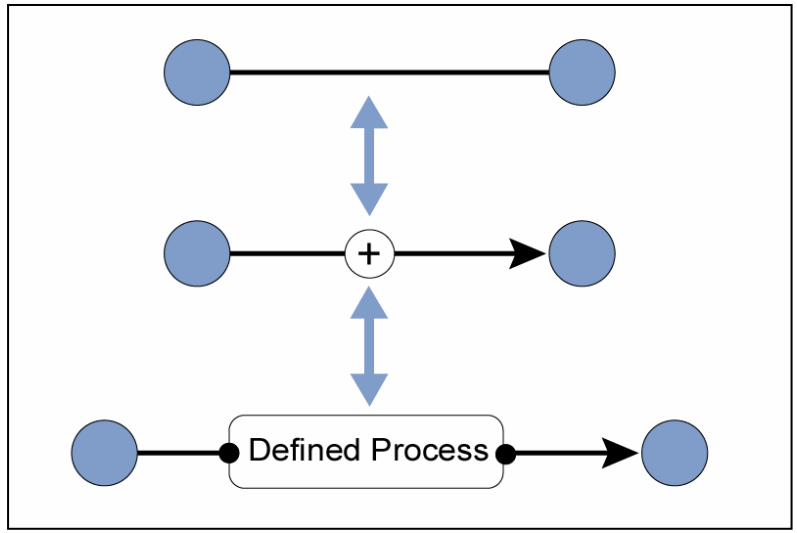


Making Connections: Multiple Players



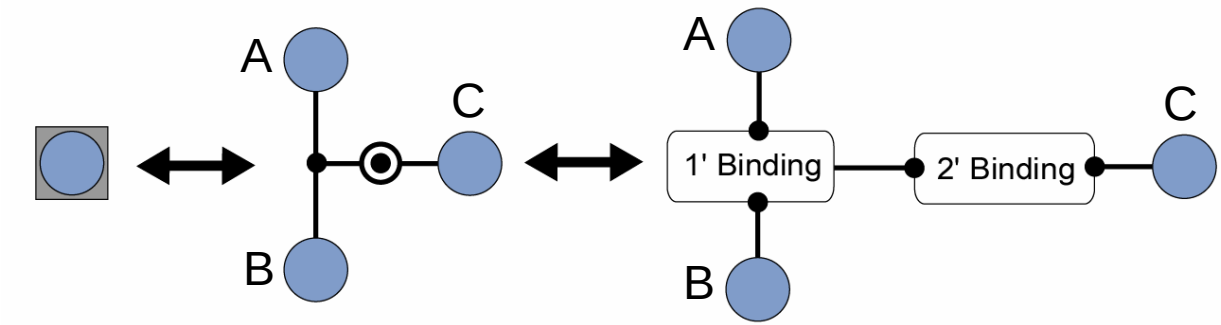
Specific interactions among molecules may require defining multiple participants. In a chemical reaction, for example, a specific regulator acts on one or more reactants to generate one or more products. Multiple participants of a reaction can be joined to a single interaction symbol to depict this process. The flow of information typically points to the product of the reaction. In the upper example, molecule A positively regulates the conversion of B to B*. Information flows from A to B*, and the role of enzyme A is indicated by the open arrow on the connecting edge. In the lower example, what is essentially the same reaction is described as an inhibition or decrease of B by A, with the consequent production of B*. Here again, flow is from A to B*. Thus, depending on the context in which a reaction is presented (i.e., the pathway being traced), it can be shown as a positive or negative interaction.

Making Connections: Defining a Specific Process



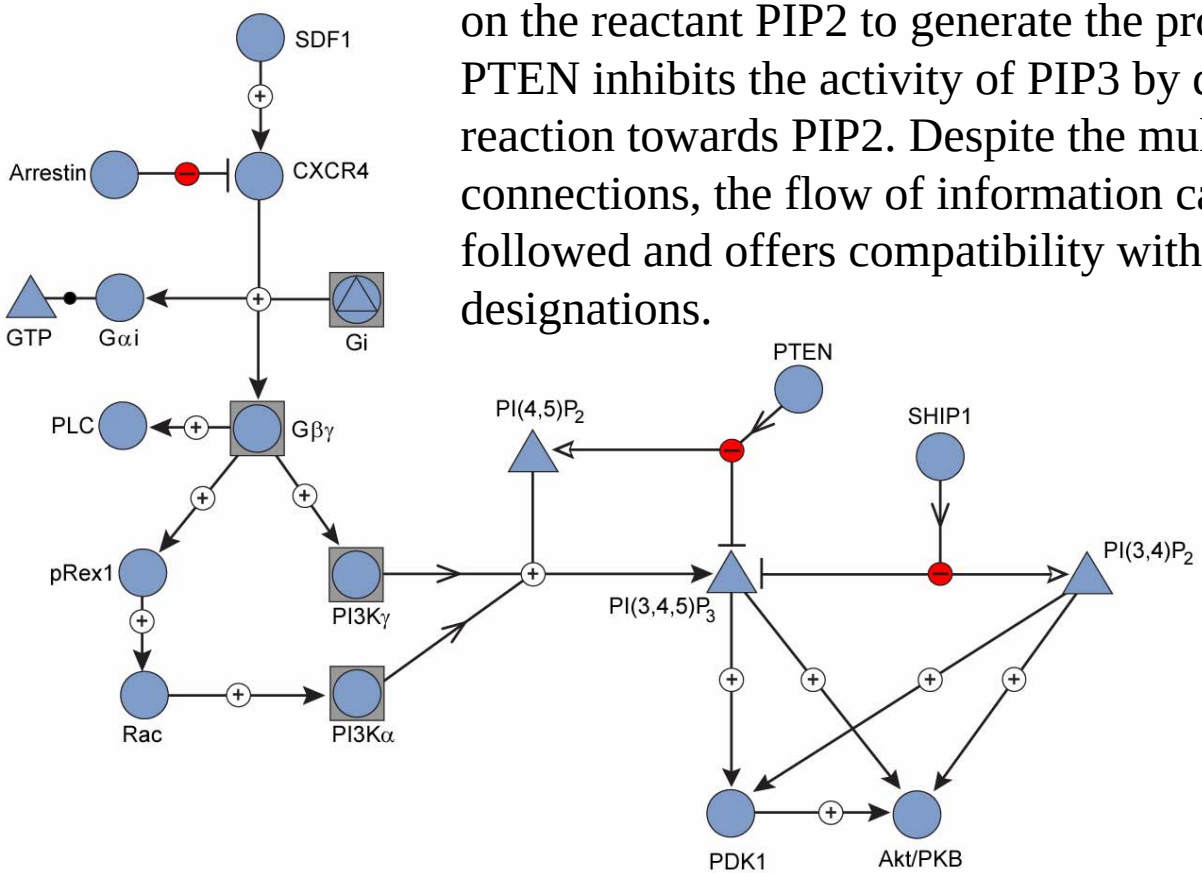
Laying out a pathway with generic positive or negative symbols may be informative enough in many cases. However, as one begins to focus in on specific paths in a signaling network, it may be useful to show what kind of specific interaction occurs among molecules. The developers of PathwayBuilder, a software program for the DARPA BioSPICE project (3), utilize a text box to define a specific process (drawn from a conceptual hierarchy of processes) that occurs among connected molecules (A. Gilman, Ph.D., University of California Berkeley, oral communication, May 2003). In the example shown, the basic interaction between A and B can be equally represented as a generic positive interaction or more specifically as a process such as phosphorylation, allosteric activation, chemical transformation, or other defined interaction or relationship. The use of an arrow as the connecting edge allows us to see the direction in which the specified process occurs. Which representation one chooses to use/view depends on the available information and on what one wishes to focus on.

Binding interactions that occur among components of a molecular complex can also be defined more specifically. For example, it might be known that molecules A and B must associate before molecule C can join the complex. As shown below, the complex created by A and B is indicated by a node along the connecting edge. An enclosed node indicates the joining of the AB complex with molecule C. Enclosed nodes can be used to indicate the joining of a complex to any molecule or other complexes. Numeric labels can also be assigned to each node to specify the binding order in more elaborate complexes. Specific information about these interactions can then be defined by expanding the complex with text boxes that identify and describe each binding interaction.



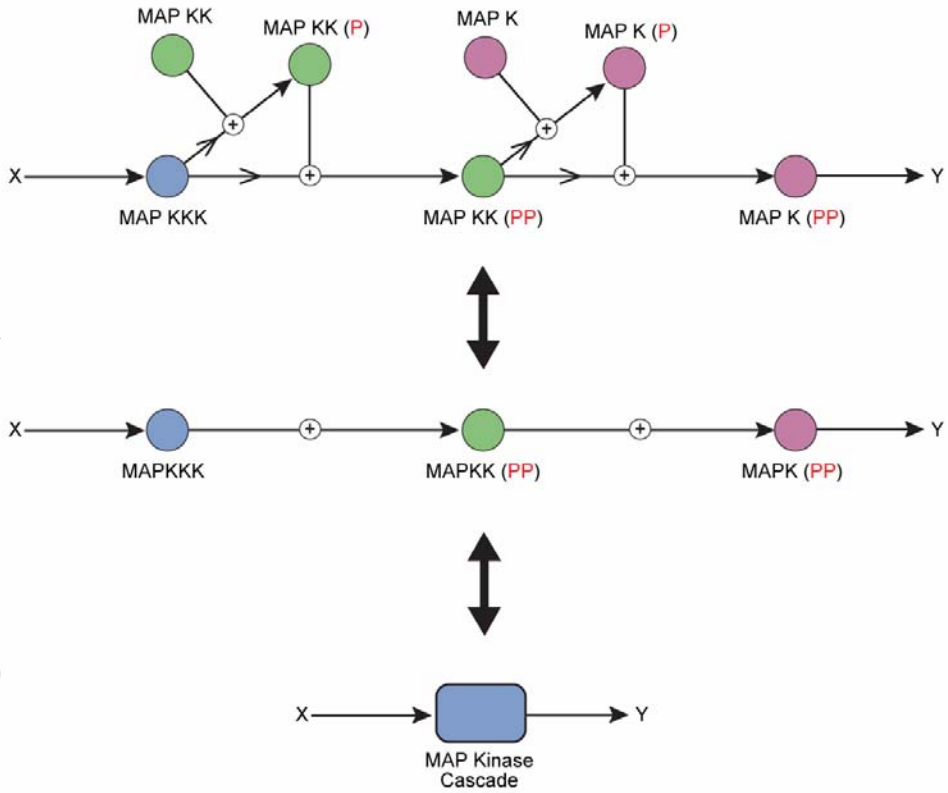
View of a Hypothetical Signaling Pathway

The hypothetical pathway depicted here utilizes the basic navigational symbols discussed. The flow of information can be easily traced from SDF1 to Akt/PKB. Note that in some places, multiple lines converge into or emerge from one interaction symbol. This is necessary when showing multiple participants of a chemical reaction. For example, the pathway shows that the PI3K isoforms act on the reactant PIP2 to generate the product PIP3. PTEN inhibits the activity of PIP3 by driving the reaction towards PIP2. Despite the multiple connections, the flow of information can be readily followed and offers compatibility with chemical designations.



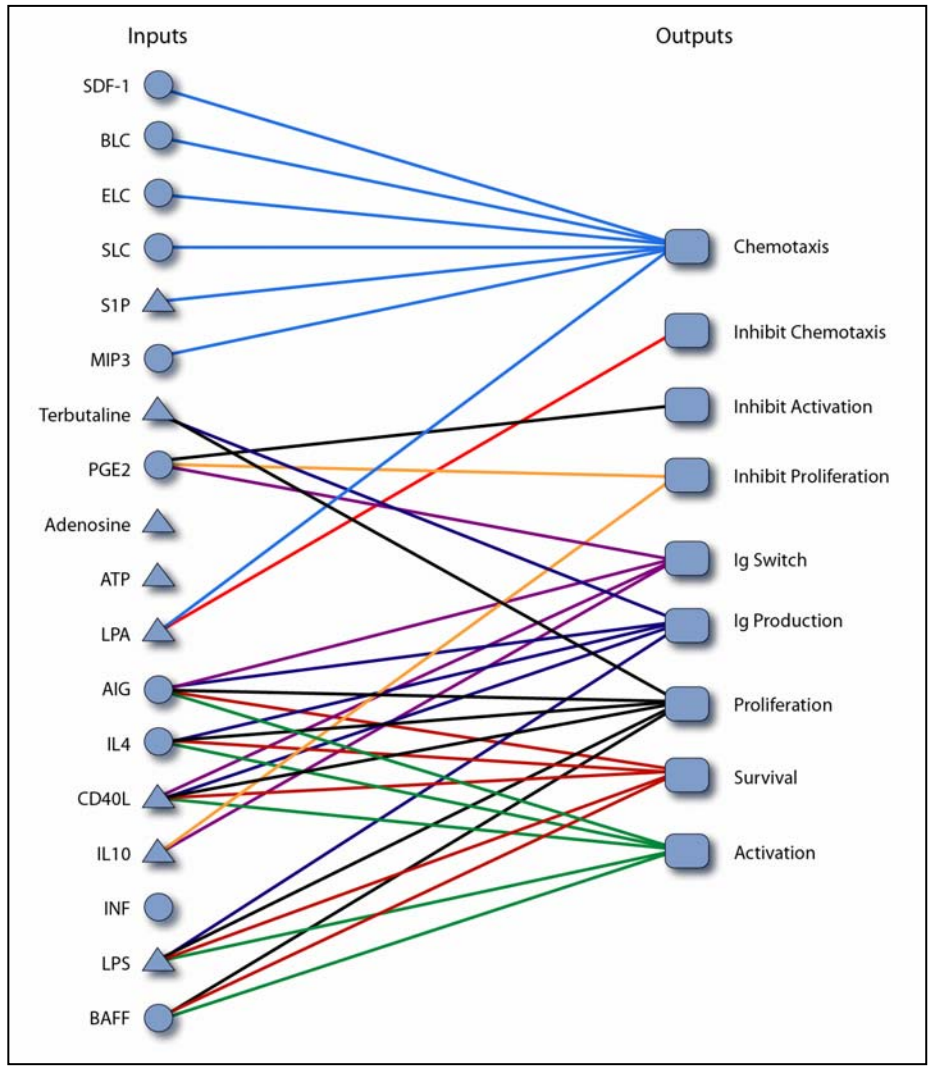
Utilizing Modules to Simplify Pathway Visualization

The MAP kinase cascade provides a good example of a commonly referenced path that can be simplified as a module(4). The flow of information from MAPKKK to MAPK requires multiple phosphorylation events. MAPKK must be phosphorylated twice in order to become active and capable of phosphorylating MAPK. MAPK must then also be dually phosphorylated to become fully active. The details of these events (shown at the top) can be collapsed into a simpler series that remains valid or packaged into a single module. Similarly, more extensive pathways can be collapsed into single modules.

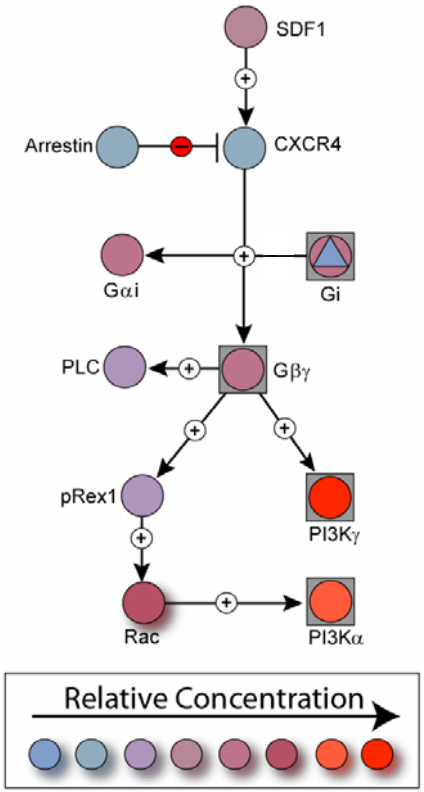


View of a Hypothetical Cellular Signaling Network

The ability to scale our field of view by representing complexes and modules as single and unique objects permits us to display a cellular network in a way that is nearly comprehensible. In this hypothetical example, all known inputs for this cell are shown with connections to all its known functional responses. All intermediates are collapsed into respective modules. This view allows us to see where paths overlap and where signaling components might be shared. This information would otherwise be hidden in a map that attempts to show all connections at once.



Color Your Path: Reveal More Information Using Different Color Keys

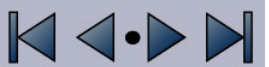


The use of color can make any pathway more informative. Different color keys can be utilized to show differences in cellular localization, enzyme activity, gene expression, relative intracellular concentrations, or any other property. In this example, we pretend to have acquired some data on relative protein concentrations for the molecules shown. The data is converted to a normalized color scale and applied to a pathway template. This adds additional viewable dimensions to the conventional pathway. Colors might also be used to highlight shared properties among signaling molecules. For example, we might wish to know which molecules in this pathway are activated by phosphorylation or which ones are translocated upon activation. Thus, various types of information about pathway components can be displayed by toggling different keys rather than constructing multiple pathway drawings.

Introduction
Conclusion
References
Contributors

Multiple Dimensions

Established pathways can be utilized to view experimental data in meaningful and dynamic ways. This idea has already been implemented in software programs such as GenMAPP, which facilitates the visualization of microarray data by grouping genes into metabolic or biochemical pathways(5). Pathways can facilitate the display of other data sets as well, including yeast two-hybrid interactions or protein expression/localization data. Algorithms that extract information of possible protein interactions from the literature have also been developed and pathways derived from this data can be very informative. Bringing these different information-rich tools together could produce a very powerful resource for signal transduction scientists. However, the union of these tools would be eased by adopting some common visual conventions whose meanings remain intact when moving from one tool to another.



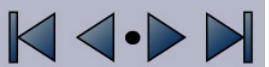
Points to Ponder

The growing complexity of signaling networks clearly poses a challenge to the visual representation of pathways. Although this communication attempts to address some of these issues, there are not always clear solutions. In any attempt to simplify a complex process, it is inevitable that some information will be lost. Here are a few unresolved issues to consider.

Time. Static maps make it difficult to appreciate the time-related dynamics of signaling processes. Tracing a path along a map is relatively easy, but viewing the temporal changes that occur in a pathway as a signal is transmitted or terminated is not.

Space. Although we like to represent signaling molecules as single entities on a map, these molecules can actually exist in multiple and changing intracellular locations at once. Intracellular localization plays an important role in signal transduction and clear ways of representing this are needed.

Context. Signaling pathways can undergo changes in their composition and subsequent outputs depending on cellular conditions. The activation of a pathway that shares components with another may affect the direction of flow for both pathways, Therefore, the reading of any map must be done with an understanding of the specific conditions being represented.



Conclusion

The organization and analysis of complex intracellular signaling networks undoubtedly requires use of computer-assisted tools and several have already been developed. But for practical reasons, experimentalists will continue to draw simple pathway maps on paper napkins and chalk boards before taking them to more sophisticated computer drafting tools. Scientists will also have a continued need to communicate their findings with simple and lucid illustrations. Consequently, signal transduction scientists require a clear language of symbols to represent pathway information, much like chemists have developed for describing chemical reactions, that can be easily transcribed from a paper to analysis software. However, the spatial and temporal dynamics of signaling pathways do require that their visual representation evolve beyond two-dimensional static wall maps.

The system of pathway navigation presented here is a work in progress meant for further exploration by interested scientists and bioinformaticists. It not only expands on some useful concepts introduced by others (1,2), but it also simplifies designations and introduces a method for scalability. Visual representations must be flexible, expandable, comprehensible, and informative. We hope this system can begin to serve that purpose. Refinement of these conventions is encouraged and it is hoped that additional contributions will lead to the development of a truly useful language.

References

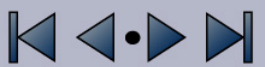
1. Kohn KW. (1999) Mol. Biol. Cell 10(8), 2703-2734. [PubMed]

2. Pirson I, Fortemaison N, Jacobs C, Dremier S, Dumont JE, and Maenhaut C. (2000) Trends Cell Biol. 10(10), 404-408. [PubMed]

3. DARPA BioSPICE Community Web site.
Available at: http://www.biospice.org

4. Ferrell JE Jr. (1996) Trends Biochem. Sci. 21(12), 460-466. [PubMed]

5. Dahlquist KD, Salomonis N, Vranizan K, Lawlor SC, and Conklin BR. (2002) Nat. Genet. 31(1), 19-20. [PubMed]



Authors*	Editors
Gilberto R. Sambrano†‡	Ashley K. Butler *Duke University, Durham, NC* Henry Bourne *University of California San Francisco, San Francisco, CA*
Acknowledgments	**Reviewers**
Adam Arkin Alex Gilman Alex Ott Shankar Subramaniam Alexandra Van Keymeulen	Adam Arkin *University of California Berkeley, Berkeley, CA* Shankar Subramaniam *University of California San Diego, La Jolla, CA*

* Please refer to the AfCS policy on authorship.
† To whom scientific correspondence should be addressed.
‡ To whom questions or comments about the *AfCS Communications* should be addressed.

Updates

2-21-2005: Figures and text on pages 3, 7, 9, 12, and 13 were updated to reflect changes in the use of open arrows () on edges. Previously, the open arrow was used to identify a substrate and also flow in an inhibitory reaction. These open arrows are now solely used to identify an enzyme or catalyst in a reaction. Closed arrows at the terminus of an edge, whether solid or white (), always point to the product of a reaction.